DSSTox Log File:

Carcinogenic Potency Database Summary Tables
(CPDBRM, CPDBHA, CPDBDG, CPDBPR)

(last updated 29 March 04)

Log of SDF Modifications and Version/revision updates:

Date	DSSTox SDF File Name	Modifications from previous version	Additional Notes
29Mar03	CPDBPR_v1b_27_15Oct03.sdf	Corrected structure DSSTox_ID=2	2,7-Acetylaminofluorene, Thanks to ACD Labs
15Oct03 15Oct03 15Oct03 15Oct03 15Oct03	CPDBRM_v1a_1354_15Oct03.sdf CPDBRM_DOP_v1a_1189_15Oct03.sdf CPDBHA_v1a_80_15Oct03.sdf CPDBDG_v1a_5_15Oct03.sdf CPDBPR_v1a_27_15Oct03.sdf	Initial launch publication; no previous versions.	Working with Source collaborators (L.S. Gold and T. H. Slone), periodic version updates to the DSSTox CPDB SDF files (i.e., v1, v2, etc.) will incorporate new information provided in updates to the CPDB Summary Tables and posted on the Source CPDB website, http://potency.berkeley.edu/. In addition, revision updates (e.g., v1a, v1b, etc) will correct reported errors or add missing data provided by users or the Source.

Description: Information in this file documents the creation, review, and update process for the DSSTox CPDB SDF files, provides summary information on database contents, and lists currently unavailable CAS information for known structures. The first section summarizes the process used for creating the initial DSSTox SDF files and the quality assurance checks and procedures employed. A table providing field and data counts offers summary overview of CPDB file contents and chemical composition. A second table provides summary counts of various types of replicate chemical information in the various CPDB files. The Log table will document any future modifications and revisions to the database content or format. To obtain the most current version of this Log File and a record of any new modifications, or to report errors in this file, a user should consult the central DSSTox website, http://www.epa.gov/nheerl/dsstox/.

QA and Development Notes:
CPDB SDF files underwent an extensive series of quality review checks prior to publication of initial launch versions. Source field entries (i.e. non-DSSTox Standard fields) were thoroughly checked by visual inspection for correspondence to original CPDB Summary Tables. We thank Lois Swirsky Gold and Thomas H. Slone for valuable assistance in ongoing quality review of the DSSTox CPDB files, helping to ensure that data are accurately extracted and represented from the original CPDB Summary Tables. They pointed out numerous systematic and human-error problems early in the DSSTox project and early in the process of CPDB SDF development, carefully reviewed DSSTox field definitions, offered suggestions for improving and finalizing all documentation files, and worked with the DSSTox team to find missing structures and reconcile remaining discrepancies in CAS numbers from the original CPDB Summary Tables.

Chemical structures were initially obtained by automated filling from large in-house databases of CAS-referenced structures (American Chemicals Directory, NCI Structure Database). The ChemFinder website (http://chemfinder.cambridgesoft.com/) was used extensively for checking CAS-to-structures and for retrieving CAS numbers for parent forms of salts and complexes. CambridgeSoft's ChemOffice 2002 ChemFinder (ver 7.0 for Windows) was used for automatic generation of SMILES codes from structures and both ChemFinder and ACD ChemFolder (ver 6.0 for Windows) were employed for "Structure-to-Name" or "Name-to-Structure" features. **ChemName**, **SMILES**, **CAS** and **Structure** field contents were checked by cross-referencing wherever possible. The CPDBRM_DOP (defined organic parent) SDF file was created by exporting only defined organics to SDF from the Main ChemFinder file for CPDBRM, and converting salts and complexes to their simplified form, with changes to corresponding Standard Chemical Fields.

Field and Data Counts in DSSTox SDF files: Refer to CPDB_FieldDefFile for definitions and explanations of all terms.

DSSTox SDF	Standard Chemical Fields	Source-specific fields	Chemical records total	Defined organic	Inorganic	Organo-metallic	Mixture or unknown*	Parent	Salt or Salt complex	Complex
CPDBRM_v1a	14	10	1354	1189	52	39	74	1016	99	165
CPDBRM_DOP_v1a	16	10	1189	1189	0	0	0	1000	67	122
CPDBHA_v1a	13	6	80	72	6	1	1	67	5	7
CPDBDG_v1a	13	4	5	5	0	0	0	4	1	0
CPDBPR_v1a	13	10	27	24	1	0	2	21	3	1

* All substances classified as **SubstanceType** = "mixture or unknown" in the CPDB data files are definitively known to be mixtures or formulations; there are no unknowns.

Replicate Information in CPDBRM SDF File: The term "replicate" refers to possibly redundant information in the chemical structure fields. All replicate cases can be easily located by search of the **ChemNote** and **ChemCount** fields in CPDBRM (refer also to CPDB_FieldDefFile).

CPDBRM: Replicate Type	Sets of Replicates	Individual Cases
CAS [1]	12	27
2D structures [2]	8	18
Parent structures [3]	30	61
Totals	50	106

[1] replicate CAS: same CAS number (e.g., if different technical grades or related mixture formulations were tested for carcinogenicity).

[2] replicate 2D structure: geometric or stereoisomers (e.g., cis and trans, RS, dl forms that might provide the same 2D information)

[3] replicate parent structures: salt or complex of same parent structure (e.g., Na and K salt of same parent structure)

Overlap of Chemical Structures or Substances in Main DSSTox CPDB SDF files:

CPDB SDF File*	CPDBRM	CPDBHA	CPDBDG	CPDBPR
CPDBRM	1354			
CPDBHA	65	80		
CPDBDG	4	0	5	
CPDBPR	24	4	0	28

* Total of 1370 distinct chemical records in 4 data files.

Wanted!! CAS Information

The listing below provides chemicals with known structures and Unknown CAS entries, which is primarily an indication of the little studied nature of these particular chemicals in the CPDB. For each, a CAS registry search was performed in CAS SciFinder and no CAS number was found by the CPDB Source authors. However, if a user has new information pertaining to any Unknown CAS in the below listing, please report this using a DSSTox Error Report Form that can be accessed from any DSSTox SDF Download Page, and be sure to indicate all relevant information (full DSSTox SDF file name, DSSTox_ID, ChemName, nature of missing information, source of correct information, etc.). Thank you!

ChemName	Structure or SMILES	CAS	SDF	Date of Request
2-Azoxypropane	[N+](=N\C(C)C)(/C(C)C)[O-]	Unknown	CPDBRM	15 Oct 03
3-Diazotyramine.HCl	C1(/C=C(\C=C/C1=O)CCN)=[N+]=[N-].[HCl]	Unknown	CPDBRM	15 Oct 03
Diethylacetylurea	O=C(N(CC)CC)NC(C)=O	Unknown	CPDBRM	15 Oct 03
Dimethylaminoethylnitrosoethylurea, nitrite salt	N(C(=O)[NH3+])(CCN=O)CCN(C)C.N(=O)[O-]	Unknown	CPDBRM	15 Oct 03
3-O-Dodecylcarbomethylascorbic acid	O(C1[C@@H]([C@@H](O)CO)OC(=O)C(O)=1)C(C(O)=O)CCCCCCCCCCCC	Unknown	CPDBRM	15 Oct 03
1-Ethylnitroso-3-(2-hydroxyethyl)-urea	O=C(N(CC)N=O)NCCO	Unknown	CPDBRM	15 Oct 03
1-Ethylnitroso-3-(2-oxopropyl)-urea	O=C(N(CC)N=O)NCC(=O)C	Unknown	CPDBRM	15 Oct 03
Hexanal methylformylhydrazone	CCCCC/C=N/N(C=O)C	Unknown	CPDBRM	15 Oct 03
1-(2-Hydroxyethyl)-nitroso-3-ethylurea	O=C(N(CCO)N=O)NCC	Unknown	CPDBRM	15 Oct 03
1-Methyl-1,4-dihydro-7-[2-(5-nitrofuryl)vinyl]-4-oxo-1,8-naphthyridine-3-carboxylate, potassium salt	[Na+].C1(=CC=C2C(=N1)N(C=C(C2=O)C([O-])=O) C)/C=C/C3=CC=C(O3)[N+]([O-])=O	Unknown	CPDBRM	15 Oct 03
3-Methylbutanal methylformylhydrazone	CC(C/C=N/N(C=O)C)C	Unknown	CPDBRM	15 Oct 03
Methylnitrosamino-N,N-dimethylethylamine	N(CCN(C)C)(C)N=O	Unknown	CPDBHA	15 Oct 03
(N-6)-(Methylnitroso)adenine	C12=C(NC=N1)N=CN=C2NCN=O	Unknown	CPDBRM	15 Oct 03
N,N-Dipropyl-4-(4'-[pyridyl-1'-oxide]azo)aniline	N(=NC1=CC=C(C=C1)N(CCC)CCC)C2=CC=[N+](C=C2)[O-]	Unknown	CPDBRM	15 Oct 03
N2-gamma-Glutamyl-p-hydrazinobenzoic acid	N(NC(CC[C@H](N)C(=O)O)=O)C1C=CC(=CC=1)C(=O)O	Unknown	CPDBRM	15 Oct 03
N-Nitrosomethyl-(2-tosyloxyethyl) amine	NC(CN=O)COS(=O)(C1=CC=C(C)C=C1)=O	Unknown	CPDBRM	15 Oct 03
N-Nitroso-ethylhydroxyethylurea	N(C(=O)N)(CCN=O)CCO	Unknown	CPDBHA	15 Oct 03
N-Nitroso-ethyl-2-oxopropylurea	N(C(=O)N)(CCN=O)CC(C)=O	Unknown	CPDBHA	15 Oct 03
N-Nitroso-oxopropylurea	N(C(=O)N)CCC(=O)N=O	Unknown	CPDBHA	15 Oct 03
N-Nitroso-oxopropylchloroethylurea	N(C(=O)N)(CCC(=O)N=O)CCCl	Unknown	CPDBHA	15 Oct 03
N-Nitroso-2-phenylethylurea	O=C(N(CCC1=CC=CC=C1)N=O)N	Unknown	CPDBHA	15 Oct 03
2-Oxopropylnitrosourea	N(C(=O)N)(N=O)CC(C)=O	Unknown	CPDBRM	15 Oct 03